

**NATIONAL SPACE SCIENCE DATA CENTER
ARCHIVE PLAN FOR 2006 - 2010**

D R A F T

Ed Grayzeck
National Space Science Data Center
Greenbelt, Maryland 20771
2006-4-25

ABSTRACT

This archive plan shows that NSSDC presently expects to accept 5-114 TB/year of space science data into the archive over the years 2006-2010.

1. INTRODUCTION

NSSDC provides a vital service as NASA's only permanent multi-disciplinary Space Science archive. Its curation activities are essential to ensure that space science data will continue to be available and usable into the indefinite future. The need for long-term curation arises because in most cases the full value of any set of data cannot be known in advance. New science discoveries or changes in research and exploration priorities may make older data, seldom noticed before, suddenly highly relevant.

This archive plan summarizes the expected data inflow to NSSDC (note the Acronym List at the end of this document), by year and by missions, for the years 2006-2010, and is the successor to several earlier plans covering 3-4 years each. Others wishing to archive data with NSSDC should contact the Head of NSSDC.

1.1 Levels of Service

NSSDC accepts and archives data under different levels of service, summarized in Table 1 below. The most familiar is the Permanent Archiving of data, but, as defined in MOUs with various data providers, it also provides Second Archive and Backup services, mostly for other Archives. The Analog Archive includes photos, maps, microfilm, microfiche, documents, etc, some analog copies of digital data and others supporting metadata.

Table 1. NSSDC Archival Storage Services	
Permanent Archive: AIPs	Preservation of digital data in Archival Information Packages delivered by a data producer or created at NSSDC. AIPs are re-written to new media within six years. Data is disseminated by NSSDC if not available through an active archive or per MOU.
Permanent Archive: non-AIP digital data	Preservation of non-packaged data on various media types. Data will eventually be migrated from legacy media to AIPs. Data is disseminated by NSSDC if not available through an active archive or per MOU.
Second Archive	Storage of digital data on distributable media that is also held by another archive. No media refreshment is performed. NSSDC may disseminate the data if authorized to do so by the primary archive as per MOU.
Backup	Storage of digital data at climate-controlled off-site facility to support another archive's contingency plan per MOU. Data will not be disseminated by NSSDC.
Analog Archive	Preservation of analog data on a variety of media with selected refreshment and selected digitization. Selected retention of original analog data after digitization. Data are copied and disseminated by NSSDC.

1.2 Archive Information Packages (AIPs)

In Table 1 NSSDC's permanent archive is digital data that is stored either as AIPs or not. The non-AIP digital data is stored on off-line media and tracked by the media on which it resides. The portion of the data stored near-line in DLT jukeboxes has been growing since 2000 and includes all new data inflows received via electronic transfer, plus some legacy data collections; it is notable not because of its media, but because those data are stored on DLTs as AIPs.

An Archive Information Package (AIP) is a single file container that holds one or many science data files, a number of attributes about each file that help NSSDC manage its AIPs, and pointers to all of the supporting documentation, including calibration information. Ideally this is enough information to allow a user to be able

to utilize the data independently of the archive and the original producer of the data. No reformatting of the science data files is performed unless record boundaries need to be retained and are not already in the byte stream. Any files that are transformed may be returned to their original state using the NSSDC defined attributes. Additionally, AIPs are media independent and platform independent.

AIPs are the preferred delivery and storage means. To that end, NSSDC makes available NSSDC packaging software (MPGA) and is encouraging Data Providers to use it to create AIPs and submit data in that form to NSSDC. The procedure assures the data submitter of the data's integrity, since the data are packaged before being sent. In the long-term most of the non-AIP data in the permanent archive is planned to be converted to AIPs.

1.3 Active Archives

NASA has established a set of Active Archives, which receive data from missions and provide electronic access to the missions' data, along with documentation and tools for accessing and using the data. NSSDC's mission is to accept data from the Active Archives or sometimes directly from missions, then provide long-term curation of the data. This is a critical service, since the full value of any set of data cannot be known in advance. New science discoveries or changes in research and exploration priorities may make older data, seldom requested, suddenly highly relevant.

2.0 ARCHIVE PLAN

The revised, detailed Archive Plan for NSSDC for 2006-2010 is given below (next page) in Table 3. Table 3 lists the missions, their launch dates and the estimated data volume to be delivered each year. Also included are the level of service (Permanent Archive, Second Archive, Backup) defined by MOU for each data collection and the discipline (Astrophysics, Heliophysics, Planetary & Lunar) for each. For archives which require Backup service, the data volumes expected from individual missions are combined and listed in the table by the name of the archive, i.e. HEASARC, IRSA, MAST, and LAMBDA.

The totals in GB for each year show an exponential growth, i.e. 5, 23, 28, 39, 107, 114 GB/yr for 2006-2010, respectively. The greatest increases are for 2009 and 2010 and are due to the Mars and Lunar Reconnaissance Orbiters (MRO & LRO). The summary of Table 3 by level of service and by discipline is given here in Table 2. Clearly, planetary missions dominate, contributing 271 TB to the NSSDC permanent archive.

TABLE 2			
Service Level	TB (2006-2010)	Discipline	TB (2006-2010)
Permanent Archive (AIPs)	271	Astrophysics	43
Second Archive	40	Heliophysics	20
Backup	6	Planetary & Lunar	254

TABLE 3. Summary of data expected at NSSDC, 2005-2010. The large increases starting in 2009 are due primarily to MRO and LRO. SDO, SOHO, Solar-B and STEREO will archive through SDAC, their data won't reach NSSDC until after 2010.

Project	Service Level * & Discipline +		Launch Date	Data Volume (GB)					
				2005	2006	2007	2008	2009	2010
ACE	A	H	1997 August	20	20	20	20	20	
AIM	A	H	2006 September			2,000	2,000	2,000	
Cassini	A	P	1997 October		2,000	2,000	4,000	11,000	
CNOFS	A	H	2006 March			1,000	1,000	1,000	
FAST	A	H	1996 August		2,000	2,000			
Genesis	A	P	2001 August	8					
Geotail	A	H	1992 July	9	1	1	1	1	
GPB	A	A	2004 April		3,000				
HEASARC	B	A		4,100	4,100	3,500	3,400	3,900	3,500
IBEX	A	H	2008 June			3	8	3	
IMAGE	A	H	2000 March	150	150	150	150	150	
IMP8	A	H	1973 October		76	2			
IRSA	B	A				5,000	2,000	2,000	2,000
ISEE 1&2	A	H	1977 October		3				
ISIS-Abuette 2	A	H	1965 November	200					
LAMBDA	B	A			350	502			810
LRO	A	P	2009 Fall					50,000	100,000
MAST	B	A			700	4,000			
Mars Express	A	P	2003 June		4,000	6,000			
Mars Odyssey	S	P	2001 April		3,000				
MER	A	P	2004 January		2,000				
MGS	S	P	1996 November	44	500				
MRO	A	P	2005 August				25,000	34,000	8,000
PDS-Other	S	P		92	50	5	348	1,700	18
Polar	A	H	1996 February		0.8	1.6	1.6	0.9	0.9
RHESSI	A	H	2002 February	1,000	1,000	1,000	1,000	1,000	
SNOE	A	H	1998 February		0.10	0.00			
THEMIS	A	H	2006 October			100	100	100	100
TIMED	A	H	2001 December		350	350			
TWINS	U	H	2006						
Ulysses	A	H	1990 October		1	1	1	1	1
Voyager	A	H	1977 August, Sept.	0.04	0.04	0.04	0.04	0.04	0.04
Wind	A	H	1994 November	2.7	2.7	2.5			
Totals (GB)				5,626	23,305	27,638	39,030	106,876	114,430

* **Service Level:** A = Permanent Archive; B = Backup; S = Second Site; U = Undetermined+ **Discipline:** A = Astrophysics; H = Heliophysics; P = Planetary & Lunar

Acronym List

ACE	Advanced Composition Explorer
AIM	Aeronomy of Ice in the Mesosphere
AIP	Archive Information Package
CNOFS	Communication/Navigation Outage Forecasting System
DLT	Digital Linear Tape
FAST	Fast Auroral Snapshot Explorer
GB	Gigabyte
GPB	Gravity Probe-B
HEASARC	High-Energy Astrophysics SARC
IBEX	Interstellar Boundary Explorer
IMAGE	Imager for Magnetopause-to-Aurora Global Exploration
IMP	Interplanetary Monitoring Platform
IRSA	Infrared Science Archive
ISEE	International Sun-Earth Explorer
ISIS	International Satellites for Ionospheric Studies
LAMBDA	Legacy Archive for Microwave Background Data Analysis
LRO	Lunar Reconnaissance Orbiter
MAST	Multi-mission Archive at Space Telescope
MER	Mars Exploration Rovers
MGS	Mars Global Surveyor
MOU	Memorandum of Understanding
MPGA	Multi-file Packager and Analyzer
MRO	Mars Reconnaissance Orbiter
NSSDC	National Space Science Data Center
PDS	Planetary Data System
RHESSI	Reuven Ramaty High Energy Solar Spectroscopic Imager
SDAC	Solar Data Archive Center
SDO	Solar Dynamics Observatory
SNOE	Student Nitric Oxide Explorer
SOHO	Solar and Heliospheric Observatory
STEREO	Solar Terrestrial Relations Observatory
TB	Terabyte
THEMIS	Time History of Events and Macroscale Interactions during Substorms
TIMED	Thermosphere Ionosphere Mesosphere Energetics and Dynamics
TWINS	Two Wide-angle Imaging Neutral-atom Spectrometers